R E P O R T R E S U M E S

ED 003 618 56

A STUDY OF THE EFFECTS OF AUTOMATED INFORMATION RETRIEVAL ON UNIVERSITY STUDENTS.

BY- PERRY, JAMES W. AND OTHERS
ARIZONA UNIV., TUCSON

REPORT NUMBER NDEA-VIIA-914 PUB DATE 3 JUN 64
GRANT OEG-7-12-0070-174-0
EDRS PRICE MF-$0.09 HC-$0.64 16P.

DESCRIPTORS- *INFORMATION RETRIEVAL, *SYSTEMS DEVELOPMENT, *SOCIAL SCIENCES, *COLLEGE STUDENTS, *STUDENT REACTION, COMPUTER BASED LABORATORIES, PILOT PROJECTS, COMPARATIVE TESTING, SOCIOLOGY, TUCSON, ARIZONA

AUTOMATED SYSTEMS WERE USED IN THIS STUDY WITH STUDENTS ENROLLED IN A LARGE UNDERGRADUATE SOCIOLOGY CLASS. THE EXPERIMENT WAS DIVIDED INTO THREE PHASES--DOCUMENTATION, EXPERIMENTATION, AND MEASUREMENT OF CRITERIA AND ANALYSES. PHASE 1 DEALT WITH THE PROBLEMS OF ABSTRACTING AND ENCODING SOCIAL SCIENCE LITERATURE AND ALSO THE DEVELOPMENT OF A SYSTEM FOR PROCESSING THE DOCUMENTS. PHASE 2 INCLUDED TESTING TO DETERMINE THE EFFECT OF INFORMATION RETRIEVAL ON 180 RANDOMLY SELECTED STUDENTS. THE SUBJECTS WERE ASSIGNED TO EITHER THE EXPERIMENTAL GROUP OR THE CONTROL GROUP. ALL SUBJECTS ATTENDED THE SAME ACTIVITIES AND RECEIVED THE SAME ASSIGNMENTS, BUT ONLY THE EXPERIMENTAL GROUP WAS PERMITTED TO USE THE AUTOMATED SYSTEM. PHASE 3 WAS DEVOTED TO CRITERIA MEASUREMENT AND ANALYSIS OF DATA. THE RESULTS INDICATED THAT NO SIGNIFICANT DIFFERENCES EXISTED BETWEEN THE GROUPS, BUT A FAVORABLE ATTITUDE EXISTED FOR USE OF THE RETRIEVAL SYSTEM. THE PILOT STUDY WAS CONTINUED FOR ANOTHER YEAR AND REVISIONS WERE MADE. THESE METHODS AND REVISIONS ARE CURRENTLY BEING WRITTEN, AND WILL BECOME AVAILABLE IN THE FUTURE. (RS)

ED003618

914

A Final Report to the

U. S. Department of Health, Education and Welfare

on

A Study of the Effects of Automated Information Retrieval on University Students,

FINAL REPORT

Grant No. 7-12-0070-174.0

Submitted by

James W. Perry
I. Roger Yoshino
Gary L. Buck

30 June 1964

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

Engineering Research Laboratory
The University of Arizona
Tucson, Arizona

Project Title: A STUDY OF THE EFFECTS OF AUTOMATED RETRIEVAL ON UNIVERSITY STUDENTS

The Problem: The problems concerned with communication through journals, books, and other documents have been accentuated during the years since World War II. The current trend emphasizing research and the investment of large sums of money in research projects indicate an even greater expansion in the ever growing fund of knowledge.

How can one cope with the problem of keeping abreast of and utilizing the vast quantity of published material? The utilization of automated systems for the storing and retrieval of recorded information offers effective means for coping with this problem.

The Project: Accordingly, under a grant from the Department of Health, Education and Welfare, a pilot study was carried out to ascertain the effects of information retrieval on students enrolled in a large undergraduate sociology class at the University of Arizona.

The first phase of the project dealt with problems of documentation related directly to the abstracting and encoding of social science literature. In order to store data in a high-speed computer for later retrieval, it was necessary to

develop a system for processing of social science documents to provide encoded standardized abstracts suitable for a computerized information. Briefly, the system for generating such abstracts entailed: (1) Defining a general standardized syntactical structure (the order to be used as means for document selection, syntactical relationships must be standardized and encoded. Such relationships were defined at the message and paragraph levels as well as at the sentence, phrase and word levels). (2) Developing a code dictionary for individual words and terms. (This involved two steps: (i) preparing a list of terms which occur most frequently in sociological literature. (ii) Encoding each term to indicate its relationship to one or more members of a carefully selected set of generic concepts or semantic factors as they may also be called.

The purpose of encoding terminology is to make both specific and generic terms available for defining searches to be performed electronically. Accordingly, in developing a list of semantic factors for machine use, the generic concepts were chosen with regard to (i) their basic importance and (ii) their intermediate generic - specific character, which may be explained as follows. On the one hand, each semantic factor was so chosen that it is not so specific as to pertain to only a few terms more specific than the factor itself. On the other hand, a semantic factor would be impaired in effectiveness if it pertains to an excessively large number of specific concepts. An example of a semantic factor is the generic concept race, which subsumes such specific terms as Negro, Caucasoid and Mongoloid.

Generation of encoded abstracts is a two step process:

1). Standardized abstracts are first prepared. (The syntactical codes* are used to indicate the general structure of the subject contents of a document. Any appropriate terminology may be used, and with exclusion of slang expressions or other terms not generally accepted. This step is performed by human analysts.)

2). The natural language terms in standardized abstracts are replaced by the corresponding semantic codes outlined above.

The development of the methods used in this research proceeded by first formulating tentative methods. Next, testing these methods by processing representative social science documents and finally, revising the tentative methods as appeared appropriate.

However, the methods were subject to continued scrutiny and revision as the generation of some 500 standardized, encoded abstracts continued. Occasionally, it appeared advisable to establish a new syntactical code in order to express an important type of syntactical relationship that had not been previously encountered. New words and terms, which were encountered as the encoding of documents continued, were assigned semantic codes and incorporated in the code dictionary. Consequently, as the number of encoded abstracts increased, the size of the code dictionary also increased but not linearly. It is interesting to note that after generating several hundred encoded abstracts, the number of times it was necessary to encode new words started

---

*Sometimes also called "role indicators," a term which has, perhaps, outlived its usefulness.

to level off. Standardized abstracts were prepared for more than 500 articles in selected social problem areas such as juvenile delinquency, minority relations and population. The ideas contained in the articles were expressed with about 2,500 encoded words, and stored in a G. E. 225 computer.

After the encoded abstracts had been recorded for electronic searching, the testing to determine the effect of information retrieval on University students was initiated. By using a table of random numbers, 180 students who were initially enrolled in a Social Problems class in the spring semester of 1963 were assigned to an experimental group and a control group. Both groups were administered the CQT scholastic aptitude test, a Concealed Figures test to measure creativity, a Cognitive Preference Test and a standardized Social Science Test. None of the above tests yielded a statistically significant difference between the control and the experimental groups. (See Table 1).

All students attended the same lecture session, took the same course examinations, and were required to write a term paper, choosing their topics from selected areas of social problems. Both groups had access to the traditional library facilities; however, only students in the experimental group were permitted the use of the information retrieval system. Students were asked to hand in their papers near the end of the semester, or approximately three months after the initial assignment. As is usually the case, after various course requirements and examinations were given, a number of students dropped out, leaving an N of 68 for the experimental group and 72 for the control group. The evaluation of the term papers was made by a five man panel; each read every paper independently and rated the papers by utilizing

Table 1

RESULTS OF PRELIMINARY TESTING

TEST	Experimental Group (N=68) Mean Score	Experimental Group (N=68) Standard Dev.	Control Group (N=72) Mean Score	Control Group (N=72) Standard Dev.
College Qualification Test (Scholastic Aptitude)	123.25	24.05	123.97	27.54
Sare-Saunders Soc. Test (Soc. Achievement)	96.32	16.36	95.71	16.52
Concealed Figures	52.21	10.79	49.64	12.95
Cognitive Preference for:				
Business Application (a)	.63	.40	.71	.75
Social Action (b)	2.35	1.50	2.07	1.27
Critical Questioning (c)	3.59	2.18	4.01	1.20
Concept Relationships (d) (Theory)	4.38	1.77	4.14	1.74

a Likert-type scale. These evaluations were then combined, across judges, into composite scores.

Did assistance from a high-speed information retrieval system provide freedom from hack work for students in the experimental group and open new opportunities for scholarship and study? Or, was the process of writing term papers made so deceptively simple that superficial and hasty efforts resulted. Was the machine process distasteful and unwanted to the highly creative students?

The third phase of the research was devoted to criteria measurement and analysis of data. Our null hypothesis was that: the introduction of machine literature searching into the normal researching process employed by students preparing term papers in an undergraduate social problems course will not:

1. affect the term paper, as measured by:
   (a) understanding of the topic
   (b) adequacy of literature search
   (c) synthesis of information
   (d) over-all quality of the paper
2. affect the attitude of the student
   (a) toward the course
   (b) toward the subject
   (c) toward the term paper assignment
3. affect the final semester grade
4. result in a different amount of incidental learning

<u>The Findings:</u>

1. The term paper grades or scores on a seven-point scale were not affected by any of the four measures of term paper quality mentioned in the hypothesis.

Table 2

RESULTS OF POST-INSTRUCTION TESTING

TEST	Experimental Group (N=68) Mean Score	Standard Dev.	Control Group (N=72) Mean Score	Standard Dev.
Attitude toward course	10.32	1.60	10.15	2.11
Attitude toward subject	11.19	1.10	11.22	1.63
Attitude toward term paper	9.81	2.89	8.10	3.91
Incidental Learning	10.82	2.82	10.85	2.69
Final Examination	100.03	17.22	100.24	18.23
Term Paper Ratings:				
Understanding of topic	4.85	.83	4.82	.91
Adequacy of search	4.49	1.16	4.35	1.19
Synthesis of information	4.60	1.01	4.49	1.08
Over-all quality	4.64	1.03	4.54	1.05

PRECEDING PAGE BLANK-NOT FILMED

Table 3

CORRELATIONS OF VARIABLES IN CONTROL AND EXPERIMENTAL GROUPS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1		602	-059	-060	-241	287	-142	013	-249	460	506	055	098	108	152	380
2	578		-063	043	-117	148	-189	034	-109	510	441	267	279	204	301	335
3	-185	-170		036	-311	-086	-007	-162	038	-152	051	-238	-257	-217	-169	-039
4	019	-174	-112		-372	-205	-303	133	108	-082	047	101	191	111	089	082
5	-102	046	-238	-477		-719	184	110	151	-145	-134	186	196	118	096	-194
6	211	154	-046	-189	-686		-194	-114	-224	257	113	-158	-206	-109	-075	202
7	-121	020	012	-090	-008	055		532	412	-250	-125	-140	-022	-122	-099	035
8	119	082	-270	-085	106	045	406		360	106	079	-000	031	017	-009	163
9	-035	007	-265	-107	111	063	299	210		-108	102	316	247	286	280	017
10	375	424	-204	011	-172	257	-049	072	-024		289	309	206	211	303	281
11	532	378	-305	-136	079	178	-011	288	013	424		287	229	315	360	190
12	258	196	-194	-011	203	-108	077	291	095	105	439		841	892	927	229
13	234	231	-160	044	114	-082	000	165	100	087	398	851		749	865	349
14	227	147	-101	075	151	-169	107	210	048	049	390	922	875		909	225
15	276	221	-187	011	164	-105	073	266	114	102	451	910	924	935		346
16	326	195	114	066	-171	110	028	072	117	209	233	123	048	151	173	

Note: Control group (N=72) correlations are entered above the diagonal; experimental group (N=68) correlations are entered below the diagonal.

1. College Qualification Test
2. Sociology pre-test score
3. Cognitive preference "a" Score
4. Cognitive preference "b" Score
5. Cognitive preference "c" Score
6. Cognitive preference "d" Score
7. Attitude toward course score
8. Attitude toward subject score
9. Attitude toward term paper
10. Incidental learning test score
11. Final exam score
12. Term paper: Understanding of topic rating
13. Term paper: Adequacy of search rating
14. Term paper: Synthesis of informating rating
15. Term paper: Over-all quality rating
16. Concealed figures test score

2. Insofar as the student's attitude toward the course and the subject matter were concerned, there was no significant difference between the experimental and the control groups. With regard to the writing of term papers, the group using the information retrieval system had a more favorable attitude. This difference was significant at the .01 level of confidence. While the time-saving as well as the psychological factors may have had a bearing, the difference was probably due to the Hawthorne effect even though precaution was taken not to introduce any bias.

3. The final semester grades, based on three multiple-choice examinations and the term paper, were comparable between the two groups.

4. There was no observable difference in incidental learning between the experimental and the control groups. (See Table 2)

Analysis of co-variance revealed that there were no substantial differences between the control and experimental groups' inter-correlation matrices. (See Table 3). Attitude scales which were administered to assess some of the by-products of the experimental condition indicated that there were no obvious negative effects. On the contrary, as it was noted, there was a favorable attitude toward the use of an automated information retrieval system.

Conclusions:

This effort serves primarily as a pilot study to ascertain the effects of information retrieval in the writing of term papers in an undergraduate sociology class. Many questions remain unanswered. For example, was the experimental group provided with enough pertinent documents in the writing of term

papers to produce statistically significant differences in the quality of the productive output of the students? Would refinement of the semantic factor list, of syntactical codes and their use in standardizing abstracts, and of the methodology for formulating queries make a difference in the experiment? While the pilot experiment did indicate that the utilization of an information retrieval system is feasible in the social sciences, it was felt that substantial refinements were in order.

Continuation of the Study through June 1964

An extensive revision of the code dictionary was made during the past year. In analyzing the dictionary, which was hurriedly constructed during the later stages of the first year of research, it became apparent that several semantic factors were grossly over used, while still others were hardly used at all.

The relative properties (RAPR, RAPS, RAPT, and RAPU) were revealed to have been used excessively and inconsistently. Another scheme of stating relative properties was devised so that relative number, intensity, distance, value, importance, etc., could be designated within a category of relative properties. This change in the dictionary necessitated a complete revision.

It further became apparent that in the first copy of the dictionary the means of stating analytical relationships (the second letter in a semantic factor) had not been exploited to full advantage. It was observed that the letters indicating analytical relationships had often been used incorrectly. This part of the revision required a consideration of each semantic factor in each code.

Particular attention was also directed to ensuring that each of the designated codes for words in the dictionary was unique. This part of the revision required a tabulation of all the numerical designators for single-factor codes up to seven-factor codes so that we could be sure that each code, although including a duplication of semantic factors, has a unique numerical representation.

After having completed these revisions, the code for each dictionary item was reconsidered and, if warranted, changed. This phase of the revision involved, as well, adding crucial words omitted from the original dictionary and deleting those words felt to be superfluous. The process of revising the dictionary cannot be finally completed as long as continuing research results in formulating new concepts and redefining old ones. A supplementary code listing, the so-called reverse code dictionary, was arranged so that each semantic factor was followed by a list of terms to which the factor in question pertained.

The revisions recently completed seem certain to facilitate further additions to the present code dictionary. Any social science discipline could use the dictionary as it now exists. The reverse code dictionary would be particularly helpful when extending the present dictionary to include additional terms especially in fields other than sociology.

The process of adding to the present dictionary would now be a relatively simple task. By consulting the reverse dictionary one could readily gain insight into the area of use of semantic factors as well as the use of such factors in combination. As noted above, the reverse dictionary provides

a listing of specific terms for each semantic factor. This listing is arranged alphabetically by each factor in each code. When using the reverse dictionary to encode new words, it would be advisable to look up one or more near synonyms in the reverse code dictionary. If a very near synonym were found, one could code the new term by merely adding a unique numerical representation to the previous code for the near synonym. But if synonymity were not close, it would be necessary extensively to modify previous code for the distant synonym or, in the extreme case, to generate an entirely new code. In any case, encoding consistency would be improved and the time and effort required would be reduced by using the previous dictionary as a guide.

Details of the methods for generating standardized abstracts, for encoding and electronically searching them and for extending the code dictionary are currently being written and will appear in book form in the near future.